



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

More than Words

Citation for published version:

Mac Sithigh, D 2010, 'More than Words: The Introduction of Internalised Domain Names and the Reform of Generic Top-level Domains at ICANN', *International Journal of Law and Information Technology*, vol. 18, no. 3, pp. 274-300. <https://doi.org/10.1093/ijlit/eqq007>

Digital Object Identifier (DOI):

[10.1093/ijlit/eqq007](https://doi.org/10.1093/ijlit/eqq007)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

International Journal of Law and Information Technology

Publisher Rights Statement:

© Mac Sithigh, D. (2010). More than Words: The Introduction of Internalised Domain Names and the Reform of Generic Top-level Domains at ICANN. *International Journal of Law and Information Technology*, 18(3), 274-300. [10.1093/ijlit/eqq007](https://doi.org/10.1093/ijlit/eqq007)

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



UEA Law School
Working Paper Series



**More than words: the introduction of internationalised
domain names and the reform of generic top-level
domains at ICANN**

Daithí Mac Síthigh

UEA LAW WPS 2010-DMS-2

ISSN 1754-5218

Other papers of the series are available at
<http://www.uea.ac.uk/law/Research/Papers>

More than words: the introduction of internationalised domain names and the reform of generic top-level domains at ICANN

Daithí Mac Síthigh^{*}

This is an electronic pre-print version of an article published in (2010) 18 *International Journal of Law & Information Technology* 274-300, available at

<http://ijlit.oxfordjournals.org/content/18/3/274>

Abstract: The Internet Corporation for Assigned Names and Numbers (ICANN) is assessed in this paper as having a special role in the development of the law of new media, recognising both the importance of its management of the global domain name system and how questions of institutional legitimacy have highlighted the lack of agreement on the role of law in the governance of the global Internet. In order to underline how ICANN's work relates to the regulation of the Internet and new media in particular, two particular issues are considered, both of which have been the subject of major announcements in 2009: (1) the facilitation of Internet multilingualism through internationalised domain names (IDNs) and (2) an attempt to expand the generic top-level domain (gTLD) system including prospect of dedicated gTLDs such as .xxx. In the case of internationalisation, it is argued that the question of internationalised domain names is best understood through its relation to historical processes of engagement between law, language and technology.

1 Introduction

The Internet Corporation for Assigned Names and Numbers (ICANN) has a task of coordinating certain aspects of the domain name system (DNS). While domain names, a familiar cultural and technical feature of the 'information society,' are a regular topic of legal scholarship, this article focuses on two points of detail within ICANN's workload, setting them in the context of the development of Internet law and earlier engagement between law and technology..

These examples have been considered appropriate for more detailed scrutiny here, with both being discussed at ICANN's various fora across a number of years and with further progress made in 2009. The first is the development of internationalised

^{*} UEA Law School, University of East Anglia, UK. Email: **@uea.ac.uk. Comments are very welcome.

domain names (IDNs), domain names using characters other than the 26 letters, 10 numbers and hyphens associated with the Roman script as used in the English language. This issue is placed in a wider context though a discussion of the regulation of multilingualism in media old and new. By approaching IDNs in this way, an argument about the relationship between law, language and technology is advanced, without relying on Internet-specific phenomena as the sole basis for scrutiny (section 2A). It is further contended that IDNs are a useful illustration of how an exercise characterised as one of technical standardisation has broader cultural consequences (section 2B). The second aspect of ICANN's work that is dealt with is the introduction of new generic top-level domains (gTLDs) (section 3).

Before turning to these issues, though, a brief consideration of the relevant technical and organisational history of the DNS and ICANN is appropriate. An Internet Protocol (IP) address (such as 123.123.4.56) identifies a server, and can be used in a Web browser to access a web page instead of the name, although few users use this approach. Naturally, the ability to definitively map names to addresses is a key element of dependability and universal accessibility. A domain name (example.com) is typically used by an end-user to access a webpage (www.example.com) or to identify a recipient of an email message (name@example.com), but must be mapped to the corresponding IP address in order to be effective. A domain name is made up of a number of segments, and reading right to left, we first find a 'top level domain' (TLD).

The stable functioning of the Internet depends on the reliability of 'root server' data, which is the modern successor to the original 'roots.txt' file that mapped the text-based domain names to the IP addresses that are the 'actual' addresses of servers connected to the Internet. The DNS is decentralised; a root server does not contain mapping data for every single website or server in the world. However, a number of key servers - the A root server and the other 12 root servers - provide information on where servers for TLDs ('generic' (gTLD) such as .com and 'country code' (ccTLD) such as .uk) are located. Each TLD administrator then maintains a server or servers containing information on servers within that namespace, and can contain

further levels of administration, such as by region (e.g. .on.ca) or by topic (e.g. .co.uk). However, the existence of a TLD from a functional point of view depends on a decision to include it in the root files. This ensures that the seeking of a website or the sending of an email is in fact reliable and possible without additional intervention. The domain name system is thus made up of more than the root servers, but they are the closest the system (and indeed the Internet) gets to a central point of control. Decisions regarding the DNS and TLDs are consequently of some importance to the wider world.

ICANN is the most visible actor in the area of domain name regulation. ICANN was born out of a complicated series of events in the late 1990s, which have been well documented by legal scholars and others.¹ It is a quasi-public consensus-driven international body, registered as a private non-profit corporation in California, and governed by a mixture of committees, advisory groups and experimental public participatory devices. Its Board exercises significant power, and has a diverse membership purporting to reflect various interests of the 'Internet community'. It does not 'register' domain names itself, but enters into agreements with registrars (entities that promote and 'sell', directly or through resellers, domain names to the end 'registrant'), or in most cases now, with managers of TLDs. In terms of the root, we should note that the US government maintains control, despite a fabled, pre-ICANN challenge to the authority of this server and the governance of the system by the then-administrators (IANA – the Internet Assigned Numbers Authority - and its manager Jon Postel), after which the US government restated and confirmed its role.² Despite proposed changes, this remained the position until 2009, with ICANN performing 'IANA functions' in connection with the global root under contract with

¹ Mueller, *Ruling the root: Internet governance and the taming of cyberspace* (MIT Press: Cambridge (MA) 2002); Paré, *Internet governance in transition: who is the master of this domain* (Rowman & Littlefield: Lanham (MD) 2003); Goldsmith & Wu, *Who Controls the Internet? : illusions of a borderless world* (OUP: New York 2006); Benoliel, 'Cyberspace Technological Standardization: an institutional theory retrospective' (2003) 18 *Berkeley Tech LJ* 1259; Klein, 'ICANN and Internet Governance: Leveraging Technical Coordination to Realize Global Public Policy' (2002) 18 *Information Society* 193; Froomkin, 'Wrong Turn In Cyberspace: Using ICANN to route around the APA and the Constitution' (2001) 50 *Duke LJ* 17; National Research Council, *Signposts in Cyberspace: The Domain Name System and Internet Navigation* (National Academies Press: Washington (DC) 2005), <http://www.nap.edu/catalog/11258.html>; and the chapters in Bygrave & Bing, *Internet Governance: infrastructure and institutions* (OUP: Oxford 2009), particularly Bing, 'Building cyberspace: a brief history of Internet'.

² Goldsmith & Wu (n 1) 44-6; Mueller (n 1) 161, 197; National Research Council (n 1).

the US Department of Commerce³ and subject to a more detailed Joint Project Agreement regarding functions and oversight. Since September 2009, ICANN enjoys a greater degree of autonomy under an ‘affirmation of commitments’ agreed with the Department of Commerce,⁴ although the formal IANA contract remains in place. ICANN is not a part of the UN system, but plays an important role in UN summits and discussions. Put simply, ICANN is legally and practically unique.

This project not only tests the evolving mechanisms for participation and lobbying that ICANN utilises, but also relates very clearly to a broader international debate on expression and censorship, while raising important questions of commerce and intellectual property. ICANN’s mission is primarily a technical one, but it has been argued - and accepted here - that its actions have broader implications within the developing international law of the Internet. Furthermore, ICANN’s unusual legal status and development of new forms of governance means that it deserves particular attention, even if the aspects discussed here are not necessarily a fair reflection of ICANN’s primary concerns and workload.

2A Law, Language and Technology

2A.1 Law and minority languages

It is first useful to consider the relationship between law, language and technology and the emphasis that is placed on the last of these by those interested in legal and linguistic issues. Crystal has argued that ‘an endangered language will progress if its speakers can make use of electronic technology’.⁵ Claims regarding the role of the Internet in language rights build upon the traditional position of seeking the

³ <http://www.iana.org/root-management.htm>.

⁴ Affirmation of commitments by the United States Department of Commerce and the Internet Corporation for Assigned Names and Numbers (30 September 2009) <http://www.icann.org/en/documents/affirmation-of-commitments-30sep09-en.htm>; see e.g. ‘ICANN be independent’ (*The Economist* 24 September 2009).

⁵ Crystal, *Language Death* (CUP: Cambridge 2000) 141, reproduced in Cunliffe & Herring, ‘Introduction to Minority Languages, Multimedia and the Web’ (2005) 11 *New Review of Hypermedia and Multimedia* 131.

availability of media of various forms in a given language.⁶ Bodies charged with the promotion of minority languages will often argue that it is necessary to ensure that media in a given language is available on the Internet. Indeed, the chief executive of Welsh-language broadcaster S4C described its online presence as comparable to the Bible being translated into Welsh,⁷ which is resonant of the discussion below on the influence of religion on the technological aspects of language policy. The role of broadcasting in the legislative and general protection of minority languages is hugely significant. Ensuring that multilingual media is available is of benefit to local language communities, to others outside of a traditional geographic heartland where one exists, and finally as a broader signal of engagement and visibility.

Multilingualism and technology is also a matter of legal concern. The European Charter for Regional or Minority Languages, for example, provides at Article 11(1) that States should encourage or facilitate broadcasting in the regional or minority language. The Charter was drawn up under the auspices of the Council of Europe (a body that has played an important role in relation to cultural and linguistic issues) in the context of the highly circumscribed competences of the European Union in this area. This again reminds us of the link between legal protection of languages and instruments of international law with limited enforcement mechanisms attached to them.; In this case, the particular value of the Charter comes from the reporting and monitoring system overseen by the Council of Europe.⁸

Furthermore, multilingualism is the subject of 'soft law' approaches through UNESCO in particular. Based on a detailed report on cultural and linguistic diversity in the 'information society',⁹ its work has included a 'multilingualism in cyberspace' action line (Initiative B@bel), reports for the World Summit on the Information Society, and even adopting 'Languages in Cyberspace' as the theme of the annual

⁶ Cunliffe & Herring, *ibid.*

⁷ Barlow, O'Malley & Mitchell, *The Media in Wales* (University of Wales Press: Cardiff 2005) 175, 153; Welsh Language Board, 'Information Technology and the Welsh Language: A Strategy Document' (2006) <http://www.byig-wlb.org.uk/English/publications/Publications/3965.pdf>

⁸ Dunbar, 'Gaelic in Scotland: the legal and institutional framework' in McLeod (ed), *Revitalising Gaelic in Scotland* (Dunedin Academic Press: Edinburgh 2006) 14-5.

⁹ UNESCO, 'Cultural and Linguistic Diversity in the Information Society' (2003) <http://unesdoc.unesco.org/images/0013/001329/132965e.pdf>

‘International Mother Language Day’ in 2006. A ‘recommendation concerning the promotion and use of multilingualism and universal access to cyberspace’¹⁰ deals with a number of substantial ‘content’ issues, and is a good example of the soft law approach to Internet regulation. It asks member states to take (‘legislative or other’) steps to secure its aims, and, in recognition of the wide range of stakeholders and bodies exercising influence on Internet content, asks them to bring the recommendation ‘to the attention of’ various bodies that can contribute to achieving its goals. In its section on access to networks and services, the recommendation includes a provision on domain name management that refers to ‘multilingual domain names’. This may seem a curious inclusion, but is in fact a reference to one of the most controversial aspects of the legal and technical aspects of international Internet regulation.

This very modern problem regarding the introduction of internationalised domain names (which will be further defined in section 2B) has raised difficult questions on the extent of language rights and the Internet. This is due to the complex relationship between language and the script in which it is written or recorded. Any language that has a written form will experience questions of script, typeface and character. It is often not an immediate concern of language law and formal legal processes, although there are some notable exceptions, such as in the constitution of India.¹¹ However, legal and other factors have historically played a role in the development of the script component of languages. One of the important other factors, as will be argued, is technology; in the present day, this is information technologies such as the Internet, WWW and email.

2A.2 Language rights and the Internet

It is appropriate to situate the problem in a broader context, that of media law. The application of media law is important in the construction of national and cultural

¹⁰ Adopted at the 32nd session of the General Conference (October 2003)

¹¹ Article 343 provides that ‘The official language of the Union shall be Hindi in Devanagari script’, while article 29.1 states that ‘Any section of the citizens residing in the territory of India or any part thereof having a distinct language, script or culture of its own shall have the right to conserve the same.’

identities. While of course we note the role of the print media (not the subject of sector-specific regulation or prior licensing in many jurisdictions), and indeed the ability of private sector players to broadcast in multiple languages, the very minority status of lesser-used languages can mean that minority language broadcasting becomes an aspect of public service broadcasting. The attention of supporters of protected languages is quickly drawn to the legislative process, as detailed legislative provisions typically regulate broadcasting at a high level (even where delivery is through an arms-length public corporation or the regulated and licensed private sector). In the case of the Internet, then, there are questions of the availability and promotion of content, but also a major technological hurdle that requires concerted international actions. There are some relatively straightforward situations where the question is merely one of the availability of content. In France, for example, the 1994 *loi Toubon* provides for the use of French in promoting or presenting a product or service, on notices in public places and in employment contracts, and is presumed to be fully capable of application to content on the Internet. Apart from one case dealing with software manuals (which involves a ‘physical’ software sale rather than a download)¹² there is little by way of firm authority on this point, although an early controversy over the website of the Georgia Tech Lorraine institution is instructive. The institution, an outpost of the (US) Georgia Institute of Technology that teaches and carries out research mostly through English in Metz, was the subject of proceedings¹³ against GTL for providing an English-only ‘advertisement’, i.e. a website. Although the case failed on procedural grounds, the institution would have argued that the *loi Toubon* was not applicable to websites, as such are ‘voluntary’ and thus ‘unlike turning on the television’.¹⁴ Nonetheless, it has been suggested that the publicity that surrounded this particular case ‘has dissuaded a large number of French companies from continuing to use English only in their websites’,¹⁵ despite some criticism from US sources, with it being argued in *Wired* that the case was a ‘bold attempt to apply

¹² *SA Surcouf* (5 February 1998), discussed in Landick, ‘French Courts and Language Legislation’ (2000) 11 *French Cultural Studies* 131, 140.

¹³ Two organisations (Défense de la langue française and Avenir de la langue française) brought the case as permitted under the law: Landick (n 12) 135.

¹⁴ Tattersall, *The Internet and the French Language* (Centre for Language in Education: Southampton 2003) 10.

¹⁵ *Ibid.*

national laws to the Internet'.¹⁶ Given the physical location of the school in French territory and the use of the .fr address, this is hardly a high water mark of Internet territorialism, particularly in the context of the approach to jurisdiction and regulation and the impact of borders on the global Internet taken by Goldsmith and Wu, who highlight the enduring ability of nation states to enforce certain provisions of national law through the control of intermediaries and of enterprises doing business within the national territory.¹⁷ Similar issues have been controversial in respect of the Charter of the French Language¹⁸ in the Canadian province of Quebec, which mentions promotional publications (s 52) software and operating systems (s 53) and advertising (s 58), all of which may affect Internet multilingualism, confirmed through cases¹⁹ and advice from the responsible authority (Office québécois de la langue française, OQLF).²⁰ Our focus here, though, will be on the impact of both law and technology on the ability to communicate in a given language, in particular where the technological aspects of a given medium facilitate or prevent the use of characters and scripts.

2A.3 Law and technology in Turkey and Japan

Languages can be regulated formally or informally, or with a mixture of approaches. The exemplar of a formal approach, of much interest to the scholar of law and information technology, is that of Turkey and Turkish. As part of the rapid social and political reforms led by the republic's founder Mustafa Kemal Atatürk in the early 20th century, Turkish went through a period of extremely rapid romanisation (conversion to the Roman script in which this article is written) backed up with unambiguous legislative force.²¹ Turkish was at the time written in Arabic/Persian

¹⁶ Ladner, 'The French Say *Non* To English-Language Web Site' (*Wired* 10 December 1996)

<http://www.wired.com/news/politics/0,1283,911,00.html>.

¹⁷ Goldsmith and Wu (n 1), especially chapters 4 and 5.

¹⁸ RSQc, C-11

¹⁹ *Procureur general du Québec v Produits métalliques CMP Ltée* (8 December 2004), cited in Racicot, 'Québec Court Upholds French Language Requirements To Website'

http://www.mccarthy.ca/pubs/publication.asp?pub_code=2099

²⁰ OQLF, 'Infoguide: Information Technologies' (June 2005), http://www.oqlf.gouv.qc.ca/english/infoguides/depliant_7_20050711.pdf. The records provided by the OQLF on its website show numerous companies have been fined for failure to comply with s 52 under the 'website' interpretation.

²¹ Lewis, *The Turkish Language Reform: a catastrophic success* (OUP: Oxford 1999) 34-8

characters, but the advocates of romanisation asserted that the ‘Turkish letters, based on the Latin’ constituted the true ‘Turkish alphabet’. Atatürk accepted the recommendations of a language commission, and a resolution to move immediately to the new script was unanimously adopted by a meeting of the Republican People’s Party in August 1928. Just a few months later, the Grand National Assembly brought this wish into force with a law on the ‘new Turkish letters’, passed on 1 November and coming into effect on two days later (with books in old script being banned in schools and forbidden to be published). Following further deadlines (including the requirement for private correspondence or records to be in the new script only) that concluded by June 1930, the matter was resolved – not even two years having passed since the Commission reported.

An alternative approach to language planning and regulation is that taken in Japan, with a mixture of formal (legal) and informal (technological) influences. The former prevailed in the early part of the century, and the latter, inevitably, became significant in recent years. The Japanese language debate, though, has also been influenced by the events of the second world war and the US occupation in 1945, and thus faced a further (and ultimately unsuccessful) threat of interventionist regulation. Japanese uses kanji (Chinese characters, ideographic) as well as hiragana and katakana (both phonetic).²² Romanisation (known in romanised Japanese as romaji) exists, but there are a number of different schema for doing so. Political groups in the early 20th century were formed to advocate for a move to either kana or romaji, and gained some support. One of their strongest arguments was the technological problems presented by a language that predominantly used kanji. For example, a kanji ‘typewriter’ was more like a printing press, and indeed the Japanese business world, let alone the home user, never really took to typewriting technology. However, the introduction of the word processor was a notable change. Users input characters on a keyboard in either kana or romaji, with the computer taking responsibility for ‘converting’ the input into kanji output.²³ This potentially undermines the pro-kana/romaji argument, as the great weakness of

²² Kanji: 日本 (Japan); Hiragana (pronunciation aid, non-kanji native simple words, etc): あえいおう (vowels A E I O U); Katakana (loanwords etc): アメリカ (America)

²³ Gottlieb, *Word processing technology in Japan: kanji and the keyboard* (Routledge: London 2000) 6, 20, 33

Japanese industry was (quite literally) changed at the stroke of a key. During the occupation by US forces at the end of the war, suggestions were made that the administration should impose romanisation through legal means as part of the reforms that were being driven through under General MacArthur's leadership. MacArthur himself, though, recognised the difficulties that this would present to good order in post-war Japan, and avoided the Atatürkian strategy.²⁴ Before and after the war, though, formal language regulation was also present. The National Language Council, which includes what would now be referred to as 'stakeholders' (educators, broadcasters, Government officials etc), is a formal body that deals with standardisation, script, etc. Japanese is unusual in that there is a Government-supported 'character list' (for kanji) – the jōyō kanji. This is legally binding for official use, and recommended for non-official use (and some newspapers have adopted it as an internal style guide). It is also used in constructing the curriculum for language learning in Japanese schools.²⁵ Significantly, though, the 'dictionaries' used for computer conversion of kana/romaji input to kanji include all the officially listed characters and more – leading to the 'rediscovery' of older kanji that users may not necessarily have prior knowledge of. The size of a dictionary has even been used as a selling point by some word processing retailers.²⁶ In both Turkish and Japanese cases, which are well documented, romanisation is a process of some controversy and difficulty, which continues across many languages and technologies to this day.

2A.4 Early information technologies and Irish language policy

In this section,²⁷ we trace the relationships between law, language and technology and the effects of such on script in the Irish language, in order to supply the necessary context for how a very similar question has caused such difficulties regarding domain names in the present day. We do this through a review of the role of law and technology in the publication of an earlier form of information technology: books and printing. By doing so, it will be possible to consider the

²⁴ See generally Unger, *Literacy and script reform in occupation Japan : reading between the lines* (OUP: Oxford 1996) 60, 75.

²⁵ Gottlieb (n 23) 106.

²⁶ Gottlieb, *Language and Society in Japan* (CUP: Cambridge 2005) 53-5

²⁷ Material in this section is drawn from files in the National Archives, Dublin, consulted by the author.

problems encountered by ICANN and Internet users as a recurring and difficult question. Irish is the first official language of Ireland, where it is spoken by 1.5 million of the 4 million residents (although fewer than a third of those speakers do so daily and the majority are in full-time education).²⁸ It is recognised under the Belfast Agreement as having a certain status in Northern Ireland and is (since 2007) an official working language of the European Union. The most recent piece of language legislation in Ireland is the Official Languages Act 2003, which provides for the use of Irish in certain contexts and areas and establishes a Commissioner to monitor compliance. Like many European languages, it uses Roman/Latin script with diacritical marks (accents, umlauts, circumflexes and other marks that modify a character). In the case of Irish, the five familiar vowels are subject to modification (lengthening) through a ‘fada’²⁹ – á é í ó ú. Consonants are also modified through lenition (séimhiú). In modern writing, this device is expressed merely with a ‘h’ after the letter (e.g. mac, mhac), but in the past, it was often represented with a dot over the consonant (mh). The use of the dot is generally understood as a part of the *cló Gaelach* (Gaelic type or script).³⁰ Other features of the Gaelic script include the use of half-uncial and miniscule styles (themselves popular in earlier centuries for Latin and other languages), and (strictly speaking) the use of the 18-letter Irish alphabet.³¹

The earliest printed books in Irish (and indeed in Gàidhlig or Scots Gaelic) were published for religious purposes. As set out in the introduction to the first book in Gàidhlig, ‘Foirm na Nurrnuideadh’ (published by the Church of Scotland in 1567), speakers of the Gaelic languages ‘labour(ed) under a disadvantage which is still greater than every other’ in not having access to religious writings in those languages.³² Subsequent publications in Irish included a (Church of Ireland) catechistical document (1571), the New Testament (1602) and a full Book of

²⁸ All statistics taken from Central Statistics Office (Ireland), *Census 2002: Volume 11 – Irish Language*, <http://www.cso.ie/census/vol11.htm>.

²⁹ The Irish word for ‘long’.

³⁰ *cló* literally translates as ‘type’, but script is often used interchangeably in English.

³¹ McGuinne, *Irish type design : a history of printing types in the Irish character* (Irish Academic Press: Dublin 1992) 1-2.

³² Ibid 163, quoting the preface written by John Carswell in *Foirm na Nurrnuideadh* (1567).

Common Prayer (1611).³³ Franciscans working in Louvain also published some books from 1611 onwards, motivated at least in part by the threat to Irish Catholics posed by the presence of non-Catholic religious books alone.³⁴ Within the Anglican tradition, there was substantial tension between those who would develop technologies (e.g. type for printing) to support Irish-language publishing for the purpose of proselytising, and those who opposed such development on the ground that it encouraged and promoted the language.³⁵ Most of these texts made use of ‘Gaelic script’ and the dot above the relevant consonants. However, a number of publications used the ‘Roman script’ used by English and other languages, with the addition of accents for vowels but not of dots. By the time of the revival in the Irish language at the end of the 19th century, there was substantial division between advocates of the ‘two scripts’.³⁶

After the establishment of the Irish Free State, language law and policy was entirely in the hands of the new authorities, who quickly provided (among other things) for the teaching of school classes through the medium of Irish.³⁷ While it is common to look at the formal provisions for language regulation in statute and constitutional law, the promotion and development of the Irish language in the area of book publishing can only be properly understood by looking at less traditional sources, as we will also do in the case of the Internet in section 2B. The role of a pro-language government in promoting a lesser-used language cannot be underestimated. For example, a substantial amount of the books published in Irish were funded through school book and general literature schemes. Therefore the standards and requirements of the commissioning Department had some influence on the language and those who used it for written expression, despite the absence of any primary or secondary legislation on this matter.

³³ United Dioceses of Dublin and Glendalough, ‘New Irish Language Version Of The Book Of Common Prayer’ (16 August 2004), <http://dublin.anglican.org/books/2004/bk040816-bcp-as-gaeilge.html> ; McGuinne (n 31) 9, 14, 18; Ó Cuív, ‘The Changing Form of the Irish Language’ in Ó Cuív (ed.), *A View of the Irish Language* (Stationery Office: Dublin 1969) 25.

³⁴ McGuinne (n 31) 24-6.

³⁵ Ibid 174.

³⁶ Ó Cuív (n 33) 26

³⁷ National Programme Conference, *Report and Programme* (Dublin: Oifig an tSoláthair, 1924).

The issue of script was a contentious one in the years following independence. Generally speaking, archive records show that the Department of Finance favoured the use of Roman script (*cló Romhánach*) on the grounds that it was a more efficient use of State resources. Typewriters, fonts for printing, and other resources were easily available in the (internationally used) Roman script than the (almost entirely Irish) *cló Gaelach*. However, the Department of Education was vigorous in defending the place of the Gaelic script in the matters under its jurisdiction, such as curricula and textbooks.. The main factor was that the teaching of Irish in schools (based, as is typical, on set books) made use of the Gaelic script. This requirement thus drove demand for books using that script. On the other hand, non-educational publishing was invariably in the Roman script, and it was argued in the 1960s that the requirement for official bilingualism (outside of education) strongly contributed to the favouring of the Roman script in the 20th century, on grounds of practicality.³⁸

One senior Minister expressed a firm view in a 1931 memorandum that type was a ‘mechanical matter on which the most competent advice could be got from a printers’ foreman’, disagreeing with the calls from various authors to be permitted to use Gaelic script or ‘old type’.³⁹ Indeed, following a change of government in 1932, the incoming Cabinet cancelled regulations prohibiting the use of Gaelic type in official communications or publications, ruling that script would be an optional matter for all concerned in future.⁴⁰ In 1963, things changed again, when the influential *Comisiún um Athbheochan na Gaeilge* (Commission on the restoration of the Irish language) called for Gaelic script to be phased out, reporting that virtually all non-educational publishing was in the Roman script, and that the accumulated

³⁸ Mac Lochlainn (National Library of Ireland) introducing a reprint of Lynam, ‘The Irish Character in Print’ (1924) 4 Transactions of the Bibliographical Society no 4; republished (Irish University Press: Shannon 1968) vii-viii.

³⁹ Minister of Finance, Eamán de Blaghd (a noted figure in the language movement). See: Handwritten note to William Doolin (Secretary, Department of Finance) dated 3 December 1928; record of meeting of the Advisory Committee (9 January 1931) quoted in a letter from the Department of Education to the Department of Finance (31 January 1931); handwritten note on the reverse of a letter from the Department of Education setting out its Advisory Committee’s motion (note dated 9 February 1931).

⁴⁰ Memorandum, 2 April 1932 (decision taken 29 March 1932).

disadvantages of Gaelic script (with particular reference to typewriting and printing) made any attempt at revival impossible and undesirable.⁴¹

Romanisation has not been the major issue for the Irish language print media that it was for other languages that make use of scripts that dramatically differ from the Roman script. In addition, given the use of Roman script in Scottish Gaelic (without question) and in Irish (in some cases, including in many official cases post-1922), the end of the Gaelic script was not as dramatic as it was in Turkish or would have been in Japanese. However, it still required much deliberation and argument, and the record shows an interesting mix of (soft) legal and administrative provisions and technological considerations. It can be argued, too, that the resolution of this long-running ‘debate’ in the 1960s could have been very different if the present-day technologies that support a diversity of scripts and characters (and buffered the pro-kanji case in Japan) had been foreseen.

The case of Irish may appear to be incapable of broader application, given the relatively small number of features that differ from roman type/Latin characters, and the clear ‘victory’ of romanisation. However, even the presence or absence of a simple diacritical mark can be, as it will be argued, a question of linguistic discrimination that is affected by both legal and technical factors. As the important moments for languages like Irish were religious translation, though, it can be appreciated how the equivalent debate in the present day relates to the availability of websites and (in this article) Web addresses in languages other than English.

2B Internationalising the DNS

2B.1 From Common Prayer to .com

Names are signifiers. Whether domain names are used as branding or for communicating the address of some particular content, the restrictions on the

⁴¹ Tuarascáil, Coimisiún um Athbheochan na Gaeilge (Dublin, 1963) [474-481] and Summary of Final Report in English (Dublin 1963) 61-3; Circular 9/63: ‘*Tá socair ag an Aire Oideachais gurb é an Cló Rómhánach a bheidh in úsáid i gceistpháipéir, i léitheoirí Gaeilge agus i dTéacsanna Próis eile ...*’.

signifier can have consequences for the signified. They are the labels of society and the study of names – and language – is generally recognised as more than a technical study of sounds and lines. A simple illustration would be the rebranding of companies in the late 1990s as ‘dot coms’ and the replacement of company names on advertisements or paraphernalia with domain names – associating the business with the perceived dynamism and success of the Internet. Studying the legal and regulatory elements of domain names, then, inevitably involves considerations of broader social relevance and a parsing of the production of meaning that goes far beyond mere considerations of interoperability and reliability..

Internationalised domain names (IDNs) are domain names that contain characters outside of the 37-character ‘LDH’ range; LDH denoting letters (A-Z), digits (0-9) and the hyphen (-). The restriction of domain name characters to LDH is something that has been present since the inception of the Internet and the DNS. Given the hierarchical structure of the DNS, any change that would expand the available characters beyond LDH must be top-down.; The link between a root server and an end user is a remote one, but the change must be sustainable at all levels, not just at the top. Internationalisation, then, cannot be brought about through decentralised autonomous actions or centralised fiat alone. The purpose of internationalisation is to facilitate multilingualism on the Internet, which as argued in section 2A is a key goal of international organisations and of various campaigners. The technical obstacles to multilingual content (as distinct from domain names) have been a matter of parallel concern, having also needed resolution for word processing, desktop publishing, data exchange, library cataloguing and so forth. The development of the Unicode standard in particular has improved the ability of Internet publishers and users to deal with a wide range of content in different languages and scripts. But the problem of IDNs stands out as an actual or perceived barrier to full multilingualism. As the President of ICANN puts it, ‘IDNs are about making the Internet more global and accessible for everyone’.⁴² The first stage of reform was the changes necessary to enable second- and subsequent-

⁴² ICANN, ‘IDN ccTLD Fast Track Process Launch’ (16 November 2009) <http://www.icann.org/en/announcements/announcement-16nov09-en.htm>

level domain names to use non-LDH characters, followed by the full implementation of IDNs through internationalised TLDs.

The technical solution to the IDN problem was discussed for many years. The ultimate solution was developed by the Internet Engineering Task Force (IETF), a loose organisation of programmers and others which has for many years set the standards that 'govern' the Internet,⁴³ and accepted by ICANN and others. The resulting 'internationalizing domain names in applications' standard (IDNA)⁴⁴ and its implementation does not involve a full restructuring of the Internet's architecture, and does not replace the DNS. Instead, it continues to use the LDH system as the basic underlying structure, while allowing for the use of IDNs using any Unicode characters that are converted (by automated process at the 'user end', such as within the web browser of an end user) to a 'Punycode' phrase that uses LDH alone, prefixed by 'xn'. An sample conversion is the name 'dail.ie' (what ICANN now refers to in the context of TLDs as an U-label), which would be represented in Punycode by xn--dil-ela.ie (A-label).

2B.2 Implementation and obstacles

The technical challenges having been dealt with, it then falls to the authorities that 'govern' the DNS to provide for the actual use of IDNs. As we have noted, the present governance systems include a complex system of committees and authorities (most of which are associated with ICANN), virtually all of which have discussed IDN implementation in detail in recent years. The various sections of ICANN's complex structure have gone on record on many occasions to express support for the introduction of IDNs. However, the process has been a long one, and is by no means complete, despite the numerous resolutions and decisions.⁴⁵ The individual managers of the country code and generic TLDs are responsible for IDN rollout in their own areas of responsibility. Some, such as.de (Germany) and

⁴³ Bowrey, *Law and Internet Cultures* (CUP: Cambridge 2005) 56ff; Alvestrand & Wium Lie, 'Development of core Internet standards: the work of IETF and W3C' in Bygrave & Bing (n 1) 126-132.

⁴⁴ National Research Council (n 1) 165-8.

⁴⁵ A collection of Board references to IDNs is collected at <http://www.icann.org/topics/idn/board-resolutions.htm>.

.org (generic) have done so, but others are still testing or have not considered the matter yet. A further issue of major difficulty was the introduction of a diversity of characters and scripts into TLDs themselves (including .com or future TLDs), which was announced with great fanfare in 2009, with a 'fast-track' process for approving individual IDN-ccTLDs announced in November 2009.⁴⁶

One of the most significant issues within the short life of IDNs to date has been that of 'spoofing' of domain names, taking advantage of the systems designed to enable IDN usage. Issues of trust in domain names have been present for many years. For example, in the early days of the popularisation and post-1995 commercialisation of the Internet, some users took advantage of the relatively informal systems for registering domain names, purchasing the names of well-known companies and derivatives thereof. This phenomenon was observed both within generic name spaces (especially .com) as well as country codes, although some of the latter had imposed restrictive policies from inception. This practice, commonly referred to as 'cybersquatting', is related to another phenomenon, inevitably termed 'typosquatting', where names similar to popular names were registered. Although the purpose of such registration varied, it was perceived as a major problem by the Internet community as well as by national authorities.⁴⁷ Purposes of 'squatting' included fraud, competition, criticism, humour and other intentions (both positive and negative). A particularly sinister example was the use of misleading web addresses (e.g. with visual confusion or an alternative location (.info instead of .com, say) to trick users into clicking on a link in an email and entering personal information that would be received (and potentially misused) by an unknown third party.

It is not surprising, therefore, that just as the expansion of the Web saw a rapid growth in cybersquatting and typosquatting, the process of IDN introduction was not devoid of what were referred to as 'homograph attacks'. The issue is quite

⁴⁶ ICANN (n 42).

⁴⁷ Clark, 'The Truth in Domain Names Act of 2003 and a preventative measure to combat typosquatting' (2004) 89 *Cornell L Rev* 1476; Moore, 'Cybersquatting: Prevention better than cure?' (2009) 17 *IJLIT* 220.

simple, and was demonstrated in early 2005 by a ‘proof of concept attack’.⁴⁸ It was demonstrated that it would be extremely simple to register a domain name using a mixture of scripts, thus leading users to be linked to a different site than the one expected. The particular example used in the demonstration was the name ‘paypal’ (the well-known electronic payments site), normally accessed through paypal.com (all characters in LDH) but also possible through paypal.com (the first a in Cyrillic script, which is visually identical to LDH ‘a’). There are other possibilities based on characters that are slightly different in alternative scripts. Although the matter was resolved, ICANN’s statement acknowledged that spoofing was a long-standing issue for Internet users, while recognising that ‘increasing the total number of characters available for domain names inevitably increases the opportunities for character confusion and spoofing’. The issue of ‘variants’ remains unresolved in the context of the proposal approval of new gTLDs using IDNs.

The systems developed to introduce IDNs, though, raise serious concerns regarding implementation that is sensitive to the right to use the language of one’s choice in relation to new media platforms. The question of spoofing discussed here, as well as practical or technical concerns and the desire to protect existing ‘names’, can lead a TLD registry to place further administrative restrictions on the possibilities presented by the technical introduction of IDNs. Are there, however, any legal arguments that might ‘nudge’ a registry in the direction of internationalisation?

2B.3 IDNs and language law in the UK, Ireland and the European Union

Of course, the international legal protection of language is weak, as discussed above. However, the individual authorities with responsibility for a TLD could potentially be the subject of control through laws of general application. Nominet, the authority responsible for .uk, has considered the introduction of IDN capabilities within its operations. At present, .uk names are restricted to LDH only. In its consultation questions, Nominet asked whether it should permit the registration of

⁴⁸ See generally ‘ICANN Statement on IDN Homograph Attacks and Request for Public Comment’ (23 February 2005), <http://www.icann.org/announcements/announcement-23feb05.htm>.

two names only distinguished by diacritic (e.g. solas, sólas or café, cafe).⁴⁹

Ultimately, the working group's report found that it should not be possible to register such a variant, although mapping of one to another should be provided for (e.g. the registrant of a name being entitled to also use the 'variant' by way of an alias).⁵⁰

This regulation, if adopted (the report was accepted in principle subject to further clarification from Nominet's executive at the Policy Advisory Board's March 2006 meeting⁵¹ but remains unimplemented) would be quite problematic. It takes what is a quite distinctive feature of many languages, and creates a linguistically and practically false distinction – that accents somehow are optional – which favours English to the exclusion of all other languages. Accented Latin characters are separate characters within languages such as Irish, and a freestanding part of the alphabet/script necessary to make correct use of that language. Therefore, providing for IDNs with one hand while preventing accurate reproduction with the other is not a good example of how to provide for online language rights. It is the extreme opposite to the *cló Gaelach* case of a unique script for a unique language – in this case, you can have any script, but it will always be inferior to that of English.

In the past, where the use of non-LDH characters was prevented by the design of the Internet, Nominet could easily be given a pass on language rights due to simple impossibility. Now, though, where the architecture is favourable to the use of languages other than English, and Nominet is voluntarily preventing their use, there is a clear policy argument against their actions. But is there also a legal one? The status of ccTLD registries varies from state to state. In the case of Nominet, it is possible that UK language law (the Welsh Language Act in particular) could be applied to Nominet actions, although this has not been established by caselaw, and Nominet's own legal advice reportedly states otherwise.⁵² Two hurdles would have

⁴⁹ http://www.nic.uk/digitalAssets/923_PabIdnConsult2Pdf.pdf (questions 5 and 6)

⁵⁰ Nominet, 'Report of the Policy Working Group on IDNs', (16 February 2006) http://www.nominet.org.uk/digitalAssets/11694_IDN_working_group_final_report.pdf.

⁵¹ Nominet, Board minutes: http://www.nic.uk/digitalAssets/6732_PAB_Report_March.pdf and http://www.nic.uk/digitalAssets/6736_PAB44.pdf.

⁵² Nominet (n 50) 2.

to be cleared - whether Nominet is a 'public body' or performs a public function,⁵³ and whether specific language legislation is indeed applicable. In a detailed study in 1999, Gould argued that Nominet could be subject to administrative law actions,⁵⁴ and although a number of aggrieved parties made threats of bringing actions against Nominet in recent years (most notably a dispute over the status of itunes.co.uk),⁵⁵ the matter remains unresolved. In the itunes.co.uk dispute, the application failed at the permission stage, based on delay and on failure to seek the available alternative remedy, with the question of Nominet's status expressly not determined.⁵⁶ While Nominet does carry with it some characteristics of a contract-based body, drawing its regulatory power from agreements between itself and registrants (which would weigh against scrutiny through public law), and has useful internal policies for dealing with disputes, its broader role is one of cultural and economic significance and resembles State action in many ways. The acquiescence of the UK state, particularly in the light of ICANN's agreements with many other states to locate ccTLD control in the national government, may engage judicial review of administrative action. Similarly, the public 'nature' of its role, where there is a realisation that the state has a part to play in Internet regulation, could engage the Human Rights Act.

In the further and narrower case of language legislation, scrutiny of the precise definition of public function or public authority used in a given case would be necessary. Indeed, the recent allocation of powers to legislate on Welsh language issues to the National Assembly for Wales⁵⁷ includes an elaborate definition of the

⁵³ The more general question of the status of a particular body for the purposes of public law (whether ordinary judicial review or the newer form of review under the Human Rights Act) has been the subject of much development in recent years. See e.g. *YL v Birmingham City Council* [2007] UKHL 27 (care homes and the Human Rights Act); *R (Weaver) v London & Quadrant Housing Trust* [2009] EWCA Civ 587 (registered social landlords and the Human Rights Act); Palmer, 'Public, private and the Human Rights Act 1998: an ideological divide' (2007) 66 *CLJ* 559

⁵⁴ Gould, 'An island In the Net: domain naming and English administrative law' (1997) 15 *John Marshall J of Computer & Info L* 493, 516

⁵⁵ McCarthy, 'Cohen disputes UK registry's legitimacy' (*The Register* 27 May 2005), http://www.theregister.ca/2005/05/27/itunes_no_nominet/; 'Nominet wins iTunes.co.uk decision' (*OutLaw News* 5 August 2005) <http://www.out-law.com/page-5979>.

⁵⁶ *CyberBritain v Nominet*, document available at http://www.nominet.org.uk/digitalAssets/30566_iTunes-Judicial-Review.pdf

⁵⁷ National Assembly for Wales (Legislative Competence) (Welsh Language) Order 2010 (SI 2010/245), under the terms of the Government of Wales Act 2006 pt 3 and sch 5.

parties that legislation may affect, including bodies outside of generally understood concepts of public functions. Indeed, other legislation may have an impact on the ability to choose a name that seems unnecessary with regard to current technology and to linguistic diversity, such as the regulation of company names.⁵⁸

In Ireland, the situation is also unclear. The .ie registry (IEDR) has had a difficult history, and is currently a non-profit private operation, having initially been maintained by University College Dublin. The Electronic Commerce Act 2000 provided (in s 31) that the relevant Minister could exercise power to regulate the 'ie domain name'. This power was subsequently transferred to ComReg (the communications regulations commission).⁵⁹ This power remains dormant, although ComReg announced in 2009 that it would adopt regulations and formally appoint IEDR as the responsible body.⁶⁰ Yet it is true to say that the influence that Government departments had over textbooks and publishing in Irish (as considered in section 2a) has not carried through to web addresses in the same language.

Under the (Irish) Official Languages Act, prescribed bodies have certain obligations with regard to the Irish language. Both the Department of Communications, Energy & Natural Resources and ComReg are prescribed for the purposes of this legislation, and thus are obliged to prepare schemes (under s 11) dealing with services provided through the Irish and English languages. Certainly, if either body was exercising the legislative power to supervise and administer .ie, it would be eminently reasonable for the question of IDNs to form part of such a scheme – or indeed, for a complaint to be made to the Commissioner regarding a failure to address the issue. If this is not the case, a slightly more tenuous argument is necessary. This would be that IEDR could be prescribed (by statutory instrument) in accordance with the provisions of the first Schedule, which allows bodies 'on which functions in relation to the general public or a class of the general public stand

⁵⁸ The Company and Business Names (Miscellaneous Provisions) Regulations 2009 (SI 1085/2009) provide (s 2 and sch 1) that the letters A-Z, the numbers 0-9 and a limited number of other marks (like £ and *) and punctuation may be used.

⁵⁹ Communications Regulation (Amendment) Act 2007 s 21.

⁶⁰ ComReg 09/01, 'Regulation of the .ie domain' (12 January 2009) http://www.comreg.ie/publications/response_to_consultation_-_regulation_of_the_ie_domain.583.103257.p.html

conferred or permitted by any enactment or by any licence or authority given under any enactment', which is quite a broad test.⁶¹ Again, we see the relevance for Internet regulation of the discussion of forms of public law as suggested in the Nominet situation.

Finally, the decision of the Court of Justice in *Konstantinidis*⁶² can be noted. Here, in the context of the civil registration of marriages, it was established that certain practices of member states in terms of transcription/transliteration could constitute a violation of what is now article 49 of the Treaty on the Functioning of the EU on freedom of establishment and non-discrimination on the grounds of nationality. As the nationality element of this principle applies to various aspects of the Treaty and its four freedoms, we can see the importance of this principle. However, the decision in this case is influenced by the impact of an obligation to use a format that modifies pronunciation and causes confusion to potential clients, and the Court argues that there is nothing in the Treaty that governs transcription per se. It is suggested that if IDN were available in general but not in a specific country, and such could be attributed to an appropriate legal provision, this case could be relied upon in an attempt to force its introduction. In this regard it is interesting to note that the EU, which has authority over the .eu ccTLD, has ensured that Greek and Cyrillic script can be used for second-level registered names.⁶³

2B.4 Observations on the impact of the IDN debates

The idea of internationalised domain names is important within the more global sphere of ICANN's work, too, as debates over script were as questions of politics and internal and external relations in decades and centuries past. For example, the Chinese government and the administrator for the .cn ccTLD, CNNIC, have been significant players in the IDN debate, through lobbying but also through some unilateral actions. Issues pertaining to alternative roots and TLDs have already been

⁶¹ Schedule 1, s 1(5)(d).

⁶² (1993) C-168/91 *Christos Konstantinidis v Stadt Altensteig*

⁶³ European Commission, '.eu' internet domain to be available also in Cyrillic and Greek alphabets' IP/09/1044 (26 June 2009)

mentioned and are also relevant in the consideration of gTLDs in section 3, below. While most alternative roots relate (without much success) to gTLDs, the Chinese attempt to develop what can be described as ‘China TLDs’ are a notable exception. Prior to the full implementation of IDNs, a number of TLDs became available within China that were regulated by CNNIC independently of the global root. These TLDs, the subject of occasional wider interest,⁶⁴ include 中国 (zhonggou - ‘China’), 公司 (gongsi - ‘company’) and 网络 (wangluo - network’). In general, it is not possible (in the absence of IDN rollout) to access a site using a ‘China TLD’ without modifying computer settings. While from the point of view of an internal user in China some argue that they are effectively second-level domains, this is still a departure from the international norm.

Root server difficulties alone are not necessarily a threat to the flow of information. However, within a jurisdiction where there is a strong effort to control Internet use, it can certainly contribute to it. This is particularly relevant in the context of ‘Chinese editions’ by non-Chinese providers such as Google that are compliant with Chinese regulations on content.⁶⁵ The desire of the Chinese government/CNNIC to have control over Chinese-language domain names, combined with the willingness to depart from the accepted root server system, indicates an uncomfortable link between recognising the importance of language within the DNS and mapping boundaries of control onto the supposedly global, borderless network. The actions of CNNIC in establishing the China TLDs can be seen in the context of the views of CNNIC’s former legal advisor, Prof. Hong Xue. Hong argued that Chinese-character domain names would be more likely to contain material causing problems with Chinese law than websites more generally, meaning that CNNIC would wish, in order to carry out content regulation, that ‘the Chinese-character domain name system be subject to Chinese administration’.⁶⁶ It is also argued from within CNNIC that ‘the US government has no right to authorize any company to manage the

⁶⁴ Mueller (n 1) 55; CNNIC Statement (3 March 2006) <http://www.cnnic.net.cn/html/Dir/2006/03/07/3623.htm>; tests by Steven Murdoch (<http://www.lightbluetouchpaper.org/2006/03/01/new-chinese-tlds/>)

⁶⁵ ‘Google censors itself for China’ (*BBC News* 25 January 2006) <http://news.bbc.co.uk/1/hi/technology/4645596.stm>.

⁶⁶ Xue, ‘The Voice of China: A story of Chinese-character domain names’ (2004) 12 *Cardozo J Int’l & Comp L* 559, 580

domain names with Chinese characters'.⁶⁷ Both approaches draw upon links between nationalism and language, but the issue encompasses a broader assertion of the role of national governments in Internet management and content regulation, and represents a difficult challenge for ICANN.

This particular situation is an example of a close mapping between language (Chinese) and state (China), which is replicated in many other cases around the world. Perhaps it is a natural and arguably inevitable response to the link between English (the only LDH language) and the legal 'home' of the DNS (the US). However, it is a worrying counterpoint not only to the pro-IDN arguments based on diversity and enabling multilingualism, but also to the belief that technology can enhance the distinctive features of a language (such as in Japanese, and theoretically for Gaelic script) in a pluralist world. The Chinese actions appropriate some of the discourse of linguistic pluralism and the internationalist/non-US position taken by many states.⁶⁸ Taken with issues of content regulation more generally, this position cannot be ignored in the eventual resolution of the IDN debate. ccTLDs have, in essence, developed into a system of international lawmaking.⁶⁹

We can better understand this point by referring to the established scholarship in sociolinguistics regarding linguistic imperialism. In particular, Skutnabb-Kangas and Phillipson (the latter having 'launched' the idea of linguistic imperialism as a subject of inquiry)⁷⁰ have analysed the international and national protection of linguistic human rights in a number of articles and texts, arguing that there are some measures that can protect individual languages but criticising the failure of international law to deal with the core problem of imperialism. A notable study for present purposes is the mapping of linguistic policy interventions in terms of overtness (overt/covert) and promotion (assimilation/maintenance) first published in

⁶⁷ Ibid 568

⁶⁸ Ermert & Hughes, 'China and the Domain Name System' in Hughes & Wacker (eds), *China and the Internet* (Routledge: London 2003) 135-6.

⁶⁹ Yu, 'The Never-Ending ccTLD Story' in Schlesinger Wass, *Addressing The World* (Rowman & Littlefield: Lanham (MD) 2003).

⁷⁰ Philipson, *Linguistic imperialism* (OUP: Oxford 1992)

1995.⁷¹ It can be added, however, that the historical regulation of the DNS is clearly situated as assimilation, although how overt it is can depend on an appreciation of the power of code as law. On the other hand, the production of multilingual content is at least based on non-discrimination (the midpoint between assimilation and maintenance), if not permission, but in a covert way due to the varying levels of legal commitment to such. The prospect of greater national control over Internet policies (including ccTLD administration) certainly indicates a greater formalisation of the regulation of language and linguistic human rights. Whether this favours assimilation or maintenance is not yet clear, and indeed may depend on the criteria under which this power is delegated/granted, and the safeguards that are built in to any such action.

The predictions that the Internet would be a vehicle for English-only content have not turned out to be entirely true. Although the fact that English is the founding and predominant language of the network and the World Wide Web cannot be ignored, the amount of non-English media content has been steadily increasing. Other languages have proven to be quite resilient, even in a general context of ‘global English’. Language laws such as those of France and Quebec can apply. Indeed, the effectively limitless capacity of the Internet (as opposed to e.g. spectrum-limited analogue TV) and the ability to access content without serious issues of time and space are factors that are potentially favourable to goals of multilingualism and language rights.

However, as seen in the discussion above, the persistence of the favouring of English within the domain name system is an uncomfortable blot on this horizon, and without correction would be an indication of a persistent linguistic imperialism. While technology has played an important role in romanisation debates in the past (whether favouring or undermining the case for such), the key role of embedded linguistic regulation with regard to the Internet is an example of the weaknesses of traditional statute-based approaches to language rights. The role of script as a

⁷¹ Skutnabb-Kangas & Phillipson, ‘Linguistic human rights, past and present’ in Skutnabb-Kangas & Phillipson (eds), *Linguistic Human Rights : Overcoming Linguistic Discrimination* (de Gruyter: Berlin 1995) 80.

feature of language – and the passions that this role has provoked – may not be a major one for all languages, but it is certainly vulnerable to use as a proxy for other issues, linguistic and cultural, assertive and defensive. Print historian Siegfried Steinberg may have argued (in the context of Irish) that it was a ‘most wholesome consequence’ of the printing press and its availability throughout the world that the Latin alphabet became ‘the one medium in which every human thought can find adequate expression’,⁷² but even if this doubtful idea was ever true to begin with, the 21st century printing press that is the Internet is (or could be) somewhat more flexible. At least a technical solution works within the current constraints of the Internet, whereas a treaty-driven restructuring of the network could well cause further problems. Governments are well experienced at regulating languages out of existence; international law has been highly deficient in protecting against this. Without strong safeguards, a well-meaning intervention to support IDNs and script diversity could narrow access to multilingual content, through the way that national authorities could use this as an opportunity to change the balance in communication and access to information.

3 Generic TLDs and public morality

3.1 Alternative and official roots

While the difficult IDN debate continued through a series of ICANN events, another TLD-related issue remained prominent. At various stages, attempts have been made to develop what ICANN now terms ‘alternative TLD name systems and roots’. Some of the earlier proponents of alternative roots, such as AlterNIC, sold then-‘unknown’ TLDs (such as ‘.ltd’ and ‘.xxx’) and ran separate root servers. For most Internet users, these services could not be used without altering software or settings).⁷³ A March 2006 report from ICANN’s security and stability advisory committee (SSAC) classified alternative roots as private, experimental, commercial,

⁷² Steinberg, *Five Hundred Years Of Printing* (Penguin: London 1955/1974) 179.

⁷³ Mueller (n 1) 130-1.

protest and political.⁷⁴ The first two are not relevant in this discussion; the final three are all outside of the root server system, for various reasons, including multilingualism. 'Universal resolution', the guarantee that a given computer connected to the Internet can access a web page or server correctly when it uses a name registered and entered in an alternative root and not in the IANA-administered system, has not been achieved.⁷⁵ Some alternative roots made arrangements for the purpose of increasing visibility (primarily commercial agreements with Internet service providers) with little success.⁷⁶ Certain governments have supported alternative root development, and may see the limited availability of servers using it as technically simple to deal with and politically advantageous, as discussed in the context of IDNs, above. However, ICANN has also (eventually) moved towards a view where expanding gTLDs (including IDN-supporting gTLDs) through more legitimate means is possible. Further work is required before these new gTLDs are approved and introduced.

3.2 Expanding gTLDs and the .xxx problem

At first, there were but seven gTLDs: .com, .org., .net being the most familiar, alongside also the more restricted .gov, .edu., .mil and .int. Adding to these TLDs was an issue throughout the popularisation of the Internet, and indeed the delays in such can be seen to have contributed to the popularity of vanity ccTLDs. In particular, those that were of appeal far beyond their territories, such as the useful .to and .tv, proposed. This has been argued to be an example of a challenge to ICANN's role but also a commodification and commercialisation of national identity,⁷⁷ and was not a sustainable approach.

The first wave of successful expansion was the introduction of a set of a further seven in 2001 (.biz, .info, .name, .pro, .coop, .aero and .museum), again divided between general interest and restricted, and a further four in 2005. It had been

⁷⁴ SSAC Report SAC009, <http://www.icann.org/committees/security/alt-tlds-roots-report-31mar06.pdf>.

⁷⁵ Ibid 15

⁷⁶ Mueller (n 1) 55.

⁷⁷ Hrynshyn, 'Globalization, nationality and commodification: the politics of the social construction of the internet' (2008) 10 *New Media & Society* 751, 763-4.

suggested that .xxx would be a useful addition to the group of gTLDs, and would have encouraged responsible activities and the clear labelling of pornographic content. An application to this effect was sponsored by a registry (ICM) in 2004 and would have been used for this genre of Web service. It was first rejected in 2005, and then, in May 2006, ICANN's board rejected the proposed gTLD. The decision was a controversial one, most notably because of the suggestions that the US government played a powerful role in the decision.⁷⁸ This theory was not confined to fringe voices, and given support by participants such as the European Union.⁷⁹ However, this is not the extent of the problems, as the proposal was – bizarrely – attacked from various directions, including by socially conservative organisations. The stage was therefore set for a further discussion and significant public controversy during 2007.

At ICANN's board meeting in Lisbon in March 2007, the application was (again) rejected. The coalition of opposing voices was even more unusual than it had been in the past, with some adult entertainment producers also arguing against the proposal, on the grounds that the registrar would wield too much power. The Board's reasons, too, were similarly confused, with arguments relating to protecting the public being joined by statements of ICANN's role as being a technical one and not related to content. In February 2010, an Independent Review Panel (IRP) published its advisory opinion on this and the previous decisions, finding that the rejection of .xxx was 'not consistent with the application of neutral, objective and fair documented policy'.⁸⁰ ICANN's Board is currently considering its options in response to this review.⁸¹

The objections to .xxx can be classified into three categories:

⁷⁸ Rogers, 'Who gives a Triple-X about Triple-X? (2007) 12 *Communications Law* 1.

⁷⁹ 'Interference by US seen in vote on .xxx domain' <http://www.nytimes.com/2006/05/11/technology/11iht-icann.html>

⁸⁰ *ICM v ICANN* (International Centre for Dispute Resolution, 19 February 2010) <http://www.icann.org/en/irp/icm-v-icann/irp-panel-declaration-19feb10-en.pdf>

⁸¹ 'ICANN Options Following the IRP Declaration on ICM's .XXX Application' <http://www.icann.org/en/irp/icm-v-icann/draft-options-post-irp-declaration-26mar10-en.pdf>

.xxx means more pornography: 'anti-porn advocates [argued] that sites would be free to keep their current '.com' address, in effect making porn more easily accessible by creating yet another channel to house it.'⁸²

This is a completely inaccurate concept of 'channel'. A website could, for example, have three different domain names for promotional purposes, perhaps in different languages, but this could not have any direct relationship with the amount of content available.

.xxx legitimises pornography.

Pornographic content is presently available on the Internet under straightforward .com names, which are at least as 'legitimate' as any other proposed name. Indeed, given the strong reputation that existing .com sites can have, the pornographic content could even be argued to gain its legitimacy from having the same gTLD as amazon.com (to take but one example), if we are to assume that having a gTLD attached to a name confers legitimacy.

.xxx would be harmful to children.

Again, this argument is distant from the reality of new media content. Whatever about plans to restrict Internet access for the purpose of child protection, the failure to approve .xxx cannot be demonstrated to have any impact on this understandably difficult issue. (There is an interpretation of the .xxx proposal that would suggest the opposite, as will be explained below).

Although there are general objections to the plan, e.g. that the primary result will be profit for the application registrar ICM,⁸³ there is also the issue highlighted by the adult entertainment industry with regard to possible censorship. This is the idea that, if .xxx were introduced, it may be used, in jurisdictions with weaker

⁸² Associated Press, 'Internet agency nixes '.xxx' Web addresses' (10 May 2006).

⁸³ Finkelstein, 'Read me first' (*The Guardian* 25 January 2007)

<http://www.guardian.co.uk/technology/2007/jan/25/comment.comment1>

constitutional protection of freedom of expression to restrict or block user access to such sites. This is far from a new idea, as various proposals for labelling or even domain-based filtering have been proposed, long before the .xxx application was a live one.⁸⁴ Examples include a consideration by a US federal commission investigating online safety⁸⁵ and various legislative proposals.⁸⁶ Academically, this has been discussed as a possible ‘zoning’ of the Internet, referring to the traditional approach under US (and other) law of regulating sexual entertainment through the zoning of land. This would depend, in part, on legislative action that would make the use of such a gTLD mandatory, as otherwise the content could be made accessible through both .xxx and another gTLD. This in itself would not necessarily be a particular legislative priority, and also relatively difficult to enforce for individual users (although admittedly easier for major players).

3.3 ICANN’s new approach

Ultimately, ICANN adopted a new policy on gTLDs in 2008, allowing a new system for applications to be discussed throughout 2009.⁸⁷ A high-profile campaign called ‘Keep the Core Neutral’⁸⁸ also took place, arguing against censorship in gTLDs. However, the ‘morality’ question was included in the process once more, with a recommendation from a sub-group (the Generic Name Supporting Organisation (GNSO)) being accepted by the Board. This recommendation was that new gTLDs ‘must not be contrary to generally accepted legal norms relating to morality and public order that are recognized under international principles of law’, with these principles including (according to the recommendation) the UDHR and other UN documents, WIPO treaties and TRIPS.⁸⁹ Adjudication would be carried out by a panel, operating independently of ICANN.

⁸⁴ See generally RFC 3675 ‘.sex Considered Dangerous’ [http:// www.ietf.org/rfc/rfc3675.txt](http://www.ietf.org/rfc/rfc3675.txt)

⁸⁵ COPA Commission, *Final Report* (20 October 2000) <http://www.copacommission.org/report/> [section E].

⁸⁶ McCullagh, ‘Senators propose labels for adult Web sites’ (*CNET News* 12 April 2007) http://news.cnet.com/2100-1028_3-6175549.html

⁸⁷ ‘Biggest Expansion in gTLDs Approved for Implementation’ (26 June 2008) <http://www.icann.org/en/announcements/announcement-4-26jun08-en.htm>

⁸⁸ <http://www.keep-the-core-neutral.org/>

⁸⁹ GNSO, ‘Introduction of new generic top-level domains’ (8 August 2007) <http://gns0.icann.org/issues/new-gtlds/pdp-dec05-fr-part-a-08aug07.htm> [Recommendation 6].

Of course, it is not unusual for a single string to be subject to morality conditions. We can take as a non-technological example the restriction on the registration of company names, which a responsible body (such as the Companies Registration Office in Ireland) must refuse if the name is ‘undesirable’.⁹⁰ This can include the protection of the public in a particular sector (such as a prohibition on the use of certain words, such as insurance or university, without the permission of relevant regulatory authorities), but also ‘offensiveness’.⁹¹ However, the troubled history of the ICANN provision, and the complex statement of purported international legal norms makes this example a little unusual. Although the system employed by ICANN is one designed for dealing with inappropriate ‘applications’, it is more than that; ICANN’s decisions are of global impact and of interest to a much wider constituency than potential gTLD applicants.

It is interesting to see how ICANN has chosen to declare international principles of law by reference to a non-exhaustive list. The inclusion of a given treaty, particularly in the light of ICANN’s unusual international legal status, should therefore be confined to ICANN’s use, but it could be interesting if there were inconsistencies (which there may well be) between ICANN’s approach and the public international law concepts of ‘general principles of law’ (ICJ Statute article 38) and ‘principles of international law’ (either in the context of Nuremberg or the ILC’s Commission on Principles of International Law). Adding further confusion, perhaps, is ICANN’s own reference to non-derogable *jus cogens*.⁹² This semi-detached relationship with public international law is characteristic of ICANN’s undefined status and the lack of progress made at multilateral institutions regarding Internet issues. This is not to suggest that a UN agency would have been able to reach a ‘better’ solution, but it is important to acknowledge the purported ability of a private corporation to use international law for its own purposes while still carrying the language and image (however confused) of existing international legal bodies. ICANN’s preliminary work on the implementation of this recommendation included consultations with scholars

⁹⁰ s 21, Companies Act 1963 (as amended by s 86 of the Company Law Enforcement Act 2001).

⁹¹ ‘Information Leaflet No. 1’ <http://www.cro.ie/en/downloads-information-leaflets.aspx>.

⁹² ICANN, ‘Morality and Public Order Objection Considerations in New gTLDs’ (draft, 29 October 2008) <http://www.icann.org/en/topics/new-gtlds/morality-public-order-draft-29oct08-en.pdf> 3.

and international judges., The most significant result was the communication of a view to ICANN by these experts that dispute resolution panels be ‘given the full scope and flexibility’ to refer to international law in deciding upon public order and morality objections. ICANN’s interim conclusion, though, is that there should be some reference to or inclusion of three key categories, namely incitement to ‘violent lawless action’, incitement to or promotion of discrimination, and incitement to or promotion of child pornography. We should note, too, the problems raised during the .xxx debate by board member Susan Crawford, who argued that ICANN continues to have difficulties with concepts of legitimacy, if it is to make decisions that are based on global Internet community values.⁹³ An appeal to international legal principles, however defined, can possibly be seen as a way of addressing this criticism in part.

4. Conclusion

The DNS was originally intended for a more humble and internal use, predating its most visible public use on the World Wide Web. The wider debate on ‘internet governance’ has frequently been consumed by arguments over domain names, and international meetings like the World Summit on the Information Society have seen lengthy debates ostensibly about domain names but engaging other issues of international and interinstitutional tension. The two particular debates discussed in this article should be considered as having confirmed ICANN’s role in the broader system of international Internet law, and the decisions in relation to explicit audiovisual and other media may foreshadow future discussions on the use of the domain name system as part of content regulation. For the time being, it is argued that it is not possible to understand the fate of a sensitive question like language in the age of information technology through traditional sources of law alone. However, this is hardly a radical observation; as discussed in section 2A, publication in Irish developed through the power of church rather than State alone, and into the 20th century, the regulation of language change and use was never a

⁹³ Crawford, ‘Why I Voted For XXX’ (30 March 2007)
http://scrawford.blogware.com/blog/_archives/2007/3/30/2845638.html

matter for legal institutions alone. By assessing the role of technological standards and limitations, and the social shaping of such technologies, a more accurate picture of language in the law of new media has been drawn. Therefore, whether for script or for language generally, the future developers and advocates of linguistic human rights can be cognisant of the context in which they operate and thus more likely to be successful in their endeavours.

The most recent developments verify the need to take ICANN seriously. In the case of internationalised domain names, the widespread discussion of this topic in the news media does point to the role of ICANN in promoting and ultimately approving an important legal and cultural shift. The BBC, for example, portrayed the 2009 decisions as a major change for the Internet as a whole (e.g. 'Internet addresses set for change'⁹⁴ and 'Web to be truly worldwide at last'⁹⁵, both published on the same day!). Although the question of gTLDs has not been answered with such clarity, the fact that ICANN requires the carrying out of extensive research and consultation with ICJ judges (and even suggesting that future decisions be made by panels made up of senior international law experts) is itself a further indication that the matter is more than an occasional decision on the approval of a name that is relevant to technical management of the Internet. The existence and public reception of both processes is a sign of the evolving role of ICANN and the way in which international legal concepts of freedom of expression, national sovereignty and multilingualism are playing a role in the management of the domain name system. There are clear consequences for the sites that will use the given domain name or TLD. Indeed, despite some concerns that were expressed at earlier stages in the processes, ICANN has managed to include a range of policy-based issues and avoided the problem faced by public authorities and States in earlier decades of focusing on technology without regard to the cultural implications of preferring or requiring one script. This is in an environment where ICANN's own legitimacy and public consultation methods remains controversial, so it is not necessarily welcome

⁹⁴ 'Internet addresses set for change' (*BBC News* 30 October 2009)

<http://news.bbc.co.uk/1/hi/technology/8333194.stm>

⁹⁵ 'Web to be truly worldwide at last' (*BBC News* 30 October 2009) <http://news.bbc.co.uk/1/hi/world/asia-pacific/8333209.stm>

that it has addressed these questions of policy and international law. The alternative, though, may have been a capitulation to a presumed (and inaccurate) technological determinism with undesirable consequences for linguistic diversity. In the case of gTLDs, it remains too soon to tell what ICANN's legacy on this point will be, and it will surely take a dispute over 'public morality' for the new process to be properly tested. How ICANN completes the two processes will be an important part of the formalisation of international Internet law with broad consequences for language and culture across the world.